Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

Entropia e Informazione in Meccanica Quantistica

Relatore:
Prof.ssa Elisa Ercolessi

Presentata da:
Eugenio Landi

Anno Accademico 2018/2019

Abstract

L'elaborato comincia con un ripasso di meccanica quantistica per poi concentrarsi sugli strumenti necessari per lo studio della teoria dell'informazione. Viene introdotto l'operatore densità e il formalismo matematico necessario per la descrizione dei sistemi aperti. L'ultimo capitolo è dedicato allo studio di quantità come entropia di Shannon e di von Neumann, viene descritto il Principio di Landauer per poi introdurre il concetto di informazione.

Indice

Introduzione

La meccanica quantistica ha avuto un enorme impatto sulla società e sulla tecnologia. Questa teoria fece la sua comparsa agli inizi del XX secolo e per rendersi conto dell'importanza che ha avuto nel mondo basta pensare ai transistor, che sono una delle più utilizzate tra le innumerevoli applicazioni della teoria quantistica. D'altro canto, un concetto altrettanto fondamentale nella nostra società è l'informazione. La grandezza fisica che misura la quantità di informazione è l'entropia e viene espressa come numero di bit necessari per immagazzinare o trasmettere l'informazione. Molte moderne discipline, come l'intelligenza artificiale, la crittografia, l'informatica e la cibernetica, si basano sulla teoria dell'informazione. Uno degli ambiti principali in cui ricopre un ruolo fondamentale riguarda lo studio dei metodi per migliorare l'affidabilità della trasmissione attraverso canali di comunicazione. Si fa uso di tecniche come la compressione dati e la correzione degli errori. Altre possibili applicazioni si trovano nella teoria dei giochi e nello studio dei mercati azionari.
L'evento decisivo che attirò l'attenzione sulla teoria dell'informazione fu la pubblicazione dell'articolo "A Mathematical Theory of Communication" da parte di Claude Shannon nel 1948. In questo articolo vengono introdotti i modelli alla base della teoria dell'informazione tramite i concetti di entropia, informazione mutua e bit. Buona parte della matematica su cui si basa la teoria dell'informazione era stata già sviluppata in precedenza nel campo della termodinamica. Il Principio di Landauer, formulato per la prima volta da Rolf Landauer nel 1961, è spesso considerato il principio base della termodinamica dell'informazione. Afferma che ad ogni operazione logica irreversibile, come l'eliminazione di un bit, è associata un aumento di entropia. Al contrario, ogni trasformazione logica reversibile dell'informazione può essere portata a termine tramite un processo termodinamicamente reversibile.
Il Principio di Landauer propone una soluzione al paradosso del '*diavoletto di Maxwell*', ma ha sollevato al contempo numerose obiezioni riguardo la sua importanza e validità.
Una conseguenza fondamentale è che, in principio, sarebbe possibile eseguire una computazione che non richiede consumo di energia semplicemente rendendo ogni step della computazione reversibile. Si può mostrare come effettivamente sia possibile svolgere qualsiasi tipo di computazione usando solo porte logiche reversibile, consentendo così una programmazione senza dispendio energetico.

In questo elaborato esporrò in un primo momento i principi su cui si basa la meccanica quantistica facendo inizialmente uso della notazione di Dirac. Verrà poi introdotto il formalismo matematico principalmente usato nella moderna teoria dell'informazione quantistica ovvero l'operatore densità, che risulta particolarmente utile nello studio dei sistemi aperti. Verranno inoltre descritte le principali proprietà che possiede e il metodo utilizzato per la descrizione degli stati puri e misti. Nel capitolo successivo verrà introdotto il concetto di superoperatore e di rappresentazione di Kraus, che consente di descrivere in modo del tutto generale i concetti di misura, evoluzione temporale e i processi stocastici che un sistema quantistico aperto può subire. L'ultimo capitolo è dedicato all'entropia e alla teoria dell'informazione mettendo in evidenza il legame con la termodinamica tramite il Principio di Landauer.

Capitolo 1

Richiami di Meccanica Quantistica

La prima sezione di questo capitolo viene dedicata ad un richiamo degli strumenti necessari per trattazione che verrà fatta. Vengono quindi esposti i quattro postulati su cui si basa la meccanica quantistica, introducendo anche il concetto delle Positive Operator Valued Measures (POVM) e viene descritto il metodo per comporre diversi sistemi quantistici. Successivamente si introduce il metodo principale di studio dei sistemi aperti: l'operatore densità, definendolo per i sistemi sia puri che misti. Ne verranno elencate le principali proprietà e verranno descritti due concetti particolarmente importanti: la decomposizione di Schmidt e la purificazione.

1.1 I postulati della teoria quantistica

La meccanica quantistica è l'ambiente matematico in cui si sviluppano le moderne teorie fisiche. Il suo sviluppo ha portato alla formulazione di quattro postulati che legano la realtà fisica al formalismo matematico usato in questa teoria.

Postulato 1: Spazio degli stati. Ad ogni sistema fisico è associato uno spazio vettoriale complesso dotato di prodotto interno (cioè uno spazio di Hilbert $\mathcal{H}$), chiamato *spazio degli stati* del sistema. Lo stato del sistema è completamente descritto dal suo *vettore di stato* appartenente a questo spazio che indichiamo con $|\psi\rangle$. Tale vettore rappresenta un *raggio* nello spazio di Hilbert del sistema. I *covettori*, o *vettori duali*, di tale spazio sono indicati con $\langle\psi|$, in accordo con la notazione di Dirac. Sono applicazioni lineari che associano ad un vettore in $\mathcal{H}$ un numero complesso $\langle\phi| : |\psi\rangle \to \langle\phi|\psi\rangle$.
Ad ogni vettore non nullo può essere associato un numero reale, detto norma, dato dal prodotto interno di un vettore con se stesso:

$$\langle\psi|\psi\rangle = ||\psi||^2. \tag{1.1}$$

È comune scegliere come rapprentante di un certo stato del sistema, un vettore di stato che abbia norma unitaria. Gli stati allora corrispondono a vettori normalizzati.

Il più semplice esempio di sistema quantistico è il *qubit*, che ha uno spazio degli stati bidimensionale. Possiamo considerare come base i due vettori ortonormali $|0\rangle$ e $|1\rangle$. Un generico vettore unitario $|\psi\rangle$ può quindi essere espresso come combinazione lineare con coefficienti complessi dei vettori di base

$$|\psi\rangle = a|0\rangle + b|1\rangle. \tag{1.2}$$

La condizione $\langle\psi|\psi\rangle = ||\psi||^2 = |a|^2 + |b|^2 = 1$ è la condizione di normalizzazione del vettore di stato. I due numeri reali $|a|^2$ e $|b|^2$ hanno la seguente interpretazione: effettuando N misurazioni sul sistema $|\psi\rangle$ si otterrà come risultato $|0\rangle$ un numero $N|a|^2$ di volte, mentre si otterrà come risultato $|1\rangle$ un numero $N|b|^2$ di volte.
Da qui è facile capire perché gli stati $e^{i\alpha}|\psi\rangle$, con α reale, e $|\psi\rangle$, che differiscono per una fase *globale*, definiscono lo stesso sistema fisico dal momento che il vettore di stato $e^{i\alpha}|\psi\rangle$ ha come vettore duale $e^{-i\alpha}\langle\psi|$ e sviluppando il prodotto le fasi complesse si semplificano.

Postulato 2: Misure. In meccanica quantistica la misura è descritta da un insieme di operatori M_m dove m indica il possibile risultato della misura. La probabilità, a priori, di misurare l'autovalore m è

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle = ||M_m|\psi\rangle||^2 \tag{1.3}$$

e il sistema, subito dopo la misura, si troverà nello stato

$$\frac{M_m|\psi\rangle}{||M_m|\psi\rangle||}. \tag{1.4}$$

Ripetendo la stessa misura si otterrà nuovamente a_n con certezza e il sistema resterà nello stesso stato. Gli operatori $\{M_m\}$ soddisfano la *relazione di completezza*

$$\sum_n M_m^\dagger M_m = I \tag{1.5}$$

che implica il fatto che $1 = \sum_m p(m)$. Consideriamo ad esempio la misura dello stato di un *qubit* sulla *base computazionale* $\{|0\rangle, |1\rangle\}$. Sono due gli esiti possibili e sono descritti dagli operatori $M_0 = |0\rangle\langle 0|$, $M_1 = |1\rangle\langle 1|$ tali che $I = M_0 + MP_1$. Supponiamo di partire da un sistema preparato nello stato $|\psi\rangle = a|0\rangle + b|1\rangle$. Allora la probabilità di ottenere 0 come risultato è

$$p(0) = \langle\psi|0\rangle\langle 0|\psi\rangle = |a|^2 \tag{1.6}$$

e lo stato del sistema dopo la misura sarà $|0\rangle$ a meno di una fase complessa senza rilevanza fisica. Analogamente, la probabilità di ottenere 1 come risultato è $p(1) = |b|^2$ e lo stato del sistema dopo la misura sarà $|1\rangle$.

Fino ad ora abbiamo postulato le misure tramite due regole: la prima ci da la probabilità di un ottenere un certo risultato e la seconda ci dice lo stato dopo aver effettuato la misura.In certi contesti lo stato del sistema dopo la misura è di poco interesse rispetto alle probabilità relativi ai diversi possibili esiti. Si pensi ad esempio al caso in cui è necessaria una sola misura per concludere l'esperimento. In queste situazioni è utile il formalismo dei POVM ('Positive Operator-Valued Measure'). Supponiamo di effettuare su un sistema $|\psi\rangle$ una misura descritta dagli operatori M_m, dove m indica il risultato della misura. La probabilità di misurare m è $p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle = \langle\psi|E_m|\psi\rangle$ dove abbiamo definito $E_m \equiv M_m^\dagger M_m$. Per la relazione di completezza vale $\sum_m E_m = I$. Gli elementi E_m sono positivi e sono noti con il nome di elementi POVM e l'insieme $\{E_m\}$ è detto POVM. Nel caso in cui $M_m = P_m$ sia ha che $E_m \equiv P_m^\dagger P_m = P_m$. È vero anche che un insieme di operatori positivi $\{E_m\}$ per cui vale la relazione di completezza forma POVM con $M_m \equiv \sqrt{E_m}$.

Il seguente esempio aiuta a chiarire l'utilità delle POVM all'interno della teoria dell'informazione quantistica. Supponiamo che Alice dia a Bob un qubit preparato in uno dei due stati

$$|\psi_1\rangle = |0\rangle, \tag{1.7}$$

oppure

$$|\psi_2\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}. \tag{1.8}$$

Come già discusso, per Bob è impossibile determinare con assoluta certezza quale stato gli è stato consegnato da Alice. Ciononostante Bob può fare un tipo di misura che consente di distinguere gli stati *alcune* volte, ma che non sbaglia *mai* identificazione. Consideriamo, infatti, la POVM descritta dai seguenti tre operatori

$$\begin{aligned} E_1 &\equiv \frac{\sqrt{2}}{1+\sqrt{2}}|1\rangle\langle 1|, \\ E_2 &\equiv \frac{\sqrt{2}}{1+\sqrt{2}}\frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2}, \\ E_3 &\equiv I - E_1 - E_2. \end{aligned} \tag{1.9}$$

Se Bob ha ricevuto lo stato $|\psi_1\rangle$ e esegue una misura descritta dalla POVM $\{E_1, E_2, E_3\}$ egli ha una probabilità nulla di misurare E_1 siccome $\langle\psi_1|E_1|\psi_1\rangle = 0$. Ciò implica che se Bob misura E_1, lo stato ricevuto da Alice era $|\psi_2\rangle$. Analogamente si può osservare che se Bob misura E_2 lo stato che ha ricevuto è $|\psi_1\rangle$. Alcune volte, però, Bob misura E_3 e non può concludere niente sull'identità dello stato che Alice gli ha passato.

Il punto chiave è che Bob non sbaglia *mai*, ma il prezzo da pagare è che, a volte, non ottiene alcuna informazione sull'identità dello stato.

Postulato 3: Evoluzione temporale. L'evoluzione temporale del vettore di stato che rappresenta un sistema fisico *chiuso* è descritta da un operatore U *unitario*, ovvero $U^\dagger U = UU^\dagger = I$, che dipende solo dagli istanti di tempo iniziale t_0 e finale t_1:

$$|\psi(t_1)\rangle = U(t_1, t_0)|\psi(t_0)\rangle. \tag{1.10}$$

A livello infinitesimo l'evoluzione del sistema è determinata dall'*equazione di Schrödinger*

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle. \tag{1.11}$$

L'operatore auto-aggiunto H è detto *Hamiltoniano* del sistema. Se quest'ultimo è indipendente dal tempo possiamo integrare a questa equazione differenziale trovando:

$$|\psi(t_1)\rangle = exp\left[\frac{-iH(t_1 - t_0)}{\hbar}\right]|\psi(t_0)\rangle \tag{1.12}$$

da cui possiamo concludere

$$U(t_1, t_0) \equiv exp\left[\frac{-iH(t_1 - t_0)}{\hbar}\right]. \tag{1.13}$$

Dal momento che H è auto-aggiunto possiamo farne la decomposizione spettrale: gli autostati di H sono chiamati autostati dell'energia e l'autovalore corrispondente rappresenta l'energia di quello stato.

Postulato 4: Osservabili. Un'osservabile è una proprietà di un sistema fisico che può essere misurata, in linea di principio. Nel contesto matematico della meccanica quantistica le osservabili sono operatori *auto-aggiunti*.

Gli operatori sono applicazioni lineari $A : |\psi\rangle \to A|\psi\rangle$ il cui aggiunto viene indicato con $A^\dagger$ ed è definito dalla relazione $\langle\psi|A^\dagger|\phi\rangle = \langle\phi|A|\psi\rangle$. Si dice auto-aggiunto un operatore A tale che $A = A^\dagger \ \ \forall|\psi\rangle,\ |\phi\rangle \in \mathcal{H}$. Dal teorema spettrale si ha che ogni operatore auto-aggiunto definito su $\mathcal{H}$ ammette una rappresentazione spettrale, ovvero i suoi autostati formano una base ortonormale, che chiameremo $|n\rangle$ e completa di $\mathcal{H}$. Possiamo dunque esprimere A in forma diagonale:

$$A = \sum_n a_n P_n \tag{1.14}$$

dove a_n sono gli autovalori di A e $P_n = |n\rangle\langle n|$ sono gli operatori di proiezione ortogonale sull'autostato $|n\rangle$ relativo all'autovalore a_n e soddisfano:

$$P_n P_m = \delta_{n,m} P_n \tag{1.15}$$

$$P_n^\dagger = P_n. \tag{1.16}$$

Il valore di aspettazione dell'osservabile A è dunque

$$\langle A\rangle \equiv \sum_n a_n p(a_n) = \sum_n a_n \langle\psi|P_n|\psi\rangle = \langle\psi|A|\psi\rangle. \tag{1.17}$$

Un altro aspetto importante della meccanica quantistica riguarda i sistemi composti. Lo spazio degli stati di un sistema composto è dato dal prodotto tensoriale degli spazi dei sistemi che lo compongono. Supponiamo di avere un sistema A con spazio degli stati $\mathcal{H}_A$ e un sistema B con spazio degli stati $\mathcal{H}_B$. Lo spazio degli stati del sistema composto AB sarà allora il prodotto tensoriale $\mathcal{H}_A \otimes \mathcal{H}_B$. Se A è preparato nello stato $|\psi_A\rangle$ e B nello stato $|\psi_B\rangle$ il sistema AB si trova nello stato $|\psi_A\rangle \otimes |\psi_B\rangle$. Date due basi ortonormali $\{|\alpha\rangle\}$ e $\{|\beta\rangle\}$, rispettivamente di $\mathcal{H}_A$ e di $\mathcal{H}_B$, allora i vettori $|\alpha,\beta\rangle \equiv |\alpha\rangle \otimes |\beta\rangle$ formano una base di $\mathcal{H}_A \otimes \mathcal{H}_B$ tale che

$$\langle\alpha,\beta|\gamma,\eta\rangle = \delta_{\alpha,\gamma}\delta_{\beta,\eta}\ . \tag{1.18}$$

1.2 L'operatore densità

Fino ad ora abbiamo formulato la meccanica quantistica con il formalismo dei vettori di stato. Una formulazione alternativa fa uso dell'*operatore densità* ρ ed è particolarmente utile nella descrizione di sistemi il cui stato non è completamente noto. Si supponga di avere un sistema che si trova in uno di un certo numero di stati $|\psi_i\rangle$, definiti su uno spazio di Hilbert $\mathcal{H}$, con probabilità p_i. Chiameremo l'insieme $\{p_i, |\psi_i\rangle\}$ *ensemble di stati puri*. È necessario, ovviamente, che

$$\sum_i p_i = 1. \tag{1.19}$$

L'operatore densità di tale ensemble è

$$\rho \equiv \sum_i p_i |\psi_i\rangle\langle\psi_i| = \sum_i p_i \rho_i. \tag{1.20}$$

Si possono riformulare i postulati della meccanica quantistica in termini dell'operatore densità mostrando come i due diversi approcci siano in realtà equivalenti.

Immaginiamo di avere un sistema come quello appena descritto. Sappiamo che la sua evoluzione temporale è governata da un operatore unitario U. Se il sistema si trova nello stato $|\psi_i\rangle$ con probabilità p_i allora, dopo l'evoluzione, si troverà nello stato $U|\psi_i\rangle$ con la stessa probabilità. Da ciò possiamo concludere che

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \longrightarrow \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^\dagger = U\rho U^\dagger. \tag{1.21}$$

Possiamo inoltre riscrivere l'equazione di Schrödinger in termini di ρ:

$$\begin{aligned}\frac{d}{dt}\rho(t) &= \left(\frac{d}{dt}|\psi(t)\rangle\right)\langle\psi(t)| + |\psi(t)\rangle\left(\frac{d}{dt}\langle\psi(t)|\right) \\ &= \frac{1}{i\hbar}H(t)|\psi(t)\rangle\langle\psi(t)| - \frac{1}{i\hbar}|\psi(t)\rangle\langle\psi(t)|H(t) \\ &= \frac{1}{i\hbar}[H(t),\rho(t)]. \end{aligned} \tag{1.22}$$

Da cui troviamo

$$i\hbar\frac{d}{dt}\rho(t) = [H(t),\rho(t)] \tag{1.23}$$

dove $[H(t),\rho(t)]$ indica il commutatore degli operatori $H(t)$ e $\rho(t)$ definito come $[H(t),\rho(t)] = H(t)\rho(t) - \rho(t)H(t)$. Misurando il valore di una osservabile A quando il sistema si trova nello stato $|\psi_i\rangle$ otteniamo il risultato a_n con probabilità

$$p(a_n|i) = \langle\psi_i|P_n|\psi_i\rangle = Tr\big[P_n|\psi_i\rangle\langle\psi_i|\big] = \sum_k \langle k|\big(P_n|\psi_i\rangle\langle\psi_i|\big)|k\rangle \tag{1.24}$$

con $|k\rangle$ base ortonormale dello spazio degli stati del sistema, dove $Tr[\ ]$ indica la *traccia* dei un operatore O. La probabilità di misurare a_n è data dalle regole della teoria della probabilità e si ha

$$p(a_n) = \sum_i p(a_n|i)p_i = \sum_i p_i Tr\big[P_n|\psi_i\rangle\langle\psi_i|\big] = Tr\big[P_n\rho\big]. \tag{1.25}$$

Per quanto riguarda lo stato dopo la misura possiamo osservare che se il sistema inizialmente si trova nello stato $|\psi_i\rangle$, allora dopo una misura che ha fornito a_n come risultato il sistema sarà nello stato

$$|\psi_i^n\rangle = \frac{P_n|\psi_i\rangle}{||P_n|\psi_i\rangle||}. \tag{1.26}$$

Perciò dopo tale misura abbiamo un ensemble di stati $\{p(i|a_n), |\psi_i^n\rangle\}$ per cui l'operatore densità sarà

$$\rho_n = \sum_i p(i|a_n)|\psi_i^n\rangle\langle\psi_i^n| = \frac{P_n\rho P_n^\dagger}{Tr\big[P_n^\dagger P_n\rho\big]}. \tag{1.27}$$

Data una base $|u_i\rangle$ di $\mathcal{H}$ possiamo esprimere ρ in componenti come

$$\langle u_i|\rho|u_j\rangle = \langle u_i|\psi\rangle\langle\psi|u_j\rangle = \psi_i{\psi_j}^* = \rho_{ij}. \tag{1.28}$$

A questo punto il valore di aspettazione di un osservabile A può essere riscritto usando le equazioni (1.25) e (1.14)

$$\langle A\rangle = \langle\psi|A|\psi\rangle = \sum_n a_n p(a_n) = \sum_n a_n Tr[P_n\rho] = Tr\big[\rho\sum_n a_n P_n\big] = Tr[A\rho]. \tag{1.29}$$

Abbiamo così mostrato che i postulati riguardanti l'evoluzione e la misura posso essere reintepretati nell'ottica dell'operatore densità.

1.3 Stati puri, misti e proprietà dell'operatore densità

Per proseguire con la nostra discussione è necessario a questo punto introdurre alcune nozioni e proprietà riguardanti l'operatore densità. Un sistema il cui stato $|\psi\rangle$ è completamente noto è detto *stato puro* e il suo operatore densità è $\rho = |\psi\rangle\langle\psi|$. Corrisponde ad un ensemble con tutte le probabilità p_i nulle, tranne una che ha valore uno. In tutti gli altri casi si ha uno *stato misto*, intendendo con ciò che tale stato è una miscela di diversi stati puri.

È possibile definire l'operatore densità prescindendo dal suo legame con i vettori di stato. Un operatore ρ è un operatore densità associato con un ensemble $\{p_i, |\psi_i\rangle\}$ se e solo se

- È auto-aggiunto: $\rho^\dagger = \rho$;
- È positivo: $\langle\psi_i|\rho|\psi_i\rangle \leq 0 \ \forall \ |\psi_i\rangle \in \mathcal{H}$;
- Ha traccia unitaria: $Tr[\rho] = 1$.

La richiesta di unitarietà della traccia rispecchia la condizione (1.19), mentre la richiesta di positività fa si che ρ ammetta una decomposizione spettrale

$$\rho = \sum_k \lambda_k |k\rangle\langle k| \tag{1.30}$$

dove $|k\rangle$ sono vettori ortogonali e λ_k sono reali, non negativi e tali che $\sum_k \lambda_k = 1$.
È possibile riformulare i postulati della meccanica quantistica tramite l'uso dell'operatore densità e ovviamente le due formulazioni sono matematicamente equivalenti.
L'operatore densità fornisce un criterio per riconoscere uno stato puro da uno misto; infatti se ρ descrive uno stato puro allora si ha $\rho^2 = \rho$, ovvero è un operatore idempotente e

$$Tr[\rho^2] = Tr[\rho] = 1. \tag{1.31}$$

Per stati misti, invece, si ha

$$\frac{1}{d} \leq Tr[\rho^2] \leq 1. \tag{1.32}$$

che $Tr[\rho^2]$ raggiunge il minimo quando gli autovalori di ρ sono tutti uguali:

$$\lambda_i = \frac{1}{d} \tag{1.33}$$

dove d è la dimensione dello spazio di Hilbert del sistema descritto da ρ.
Vi è inoltre un certo grado di libertà nell'ensemble che definisce una certa matrice densità. Due diversi insiemi di vettori $|\psi_i'\rangle$ e $|\phi_j'\rangle$, non necessariamente normalizzati, *generano* lo stesso operatore densità se e solo se

$$|\psi_i'\rangle = \sum_j u_{ij} |\phi_j'\rangle \tag{1.34}$$

con u_{ij} matrice unitaria di numeri complessi.

1.4 Sistemi composti, operatore densità ridotto

L'applicazione più interessante dell'operatore densità riguarda i *sottosistemi*. Immaginiamo di avere un sistema AB formato dai due sottosistemi A e B. Se il sistema si trova in uno stato puro vale $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ e possiamo scrivere

$$|\psi\rangle = \sum_{i,j} c_{ij} |\alpha_i\rangle \otimes |\beta_j\rangle \tag{1.35}$$

dove $|\alpha_i\rangle$ e $|\beta_j\rangle$ sono basi ortonormali complete, rispettivamente di $\mathcal{H}_A$ e $\mathcal{H}_B$, mentre c_{ij} sono coefficienti complessi normalizzati tali che

$$\sum_{i,j} |c_{ij}|^2 = 1. \tag{1.36}$$

Consideriamo un operatore $O = O_A \otimes I_B$ che agisce solo sul sottosistema A, lasciando il sottosistema B inalterato. Il valore di aspettazione di tale osservabile è:

$$\begin{aligned}
\langle\psi|O_A \otimes I_B|\psi\rangle &= \Big(\sum_{i,j} c_{ij}^* \langle\alpha_i| \otimes \langle\beta_j|\Big) O_A \otimes I_B \Big(\sum_{k,l} c_{kl} |\alpha_k\rangle \otimes |\beta_l\rangle\Big) \\
&= \sum_{i,j,k,l} c_{ij}^* c_{kl} \langle\alpha_i|O_A|\alpha_k\rangle \langle\beta_j|\beta_l\rangle \\
&= \sum_{i,j,k,l} c_{ij}^* c_{kl} \langle\alpha_i|O_A|\alpha_k\rangle \delta_{jl} \\
&= \sum_{i,k} \Big(\sum_j c_{ij}^* c_{kj}\Big) \langle\alpha_i|O_A|\alpha_k\rangle \\
&= \sum_i \langle\alpha_i|\rho_A O_A|\alpha_i\rangle.
\end{aligned} \tag{1.37}$$

dove abbiamo usato l'ortonormalità della base $|\beta_j\rangle$ e abbiamo definito l'operatore densità ridotto

$$\rho^A \equiv Tr_B[\rho^{AB}] = \sum_n \langle\beta_n|\big(|\psi\rangle\langle\psi|\big)|\beta_n\rangle \tag{1.38}$$

i cui elementi matriciali rispetto alla base $|\alpha_i\rangle$ sono

$$\langle\alpha_i|\rho^A|\alpha_k\rangle = \sum_j c_{ij}^* c_{kj} \ . \tag{1.39}$$

Tr_B è una mappa tra operatori chiamata *traccia parziale* sul sistema B. Si può dimostrare che la traccia parziale è l'*unica* operazione che dà origine alla corretta descrizione delle quantità osservabili per sottosistemi di sistemi compositi. La matrice densità ridotta permette di studiare il comportamento di un solo sottosistema, tenendo conto dell'altro sottosistema e della interazione tra questi due.

L'operatore densità ridotto ha le stesse proprietà dell'operatore densità:

- È auto-aggiunto: $\rho^{A^\dagger} = \rho^A$;
- È positivo: $\langle\phi|\rho^A|\phi\rangle \leq 0 \ \forall \ |\phi\rangle \in \mathcal{H}_A$;
- Ha traccia unitaria: $Tr[\rho^A] = 1$.

Un esempio particolarmente interessante è lo stato di Bell:

$$|\psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \ . \tag{1.40}$$

A questo stato è associato il seguente operatore di densità:

$$\begin{aligned} \rho &= |\psi\rangle\langle\psi| \\ &= \frac{|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|}{2} \ . \end{aligned} \tag{1.41}$$

Se facciamo la traccia parziale sul secondo qubit, otteniamo l'operatore di densità ridotto del primo qubit:

$$\rho^1 = Tr_2[\rho] = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} = \frac{I}{2} \ . \tag{1.42}$$

Possiamo notare che questo stato è uno stato misto, dal momento che:

$$Tr\left[\left(\frac{I}{2}\right)^2\right] = \frac{1}{2} < 1 \tag{1.43}$$

Lo stato del sistema composto dei due qubit è uno stato puro, tuttavia il primo qubit si trova in uno stato misto.

1.5 Decomposizione di Schmidt e purificazione

Per proseguire la discussione e poter parlare della Teoria dell'Informazione Quantistica è opportuno introdurre due strumenti aggiuntivi di grande utilità: la *decomposizione di Schmidt* e la *purificazone*.

Come in (1.35), possiamo esprimere un vettore arbitrario di $\mathcal{H}_A \otimes \mathcal{H}_B$ come

$$|\psi\rangle = \sum_{i,j} c_{ij}|\alpha_i\rangle|\beta_j\rangle \equiv \sum_i |\alpha_i\rangle|\widetilde{\beta}_i\rangle \tag{1.44}$$

dove $|\widetilde{\beta}_i\rangle = \sum_j c_{ij}|\beta_j\rangle$ possono non essere ortonormali. Supponendo che i ket $|\alpha_i\rangle$ formano una base ortonormale che diagonalizza ρ^A con autovalori $\{p_i\}$, allora la matrice densità ridotta è data da

$$\rho^A = \sum_i p_i|\alpha_i\rangle\langle\alpha_i| \ . \tag{1.45}$$

Possiamo ricavare ρ^A tramite la traccia parziale

$$\rho^A = Tr_B\big[|\psi\rangle\langle\psi|\big] = Tr_B\bigg[\sum_{i,j} |\alpha_i\rangle\langle\alpha_j| \otimes |\widetilde{\beta}_i\rangle\langle\widetilde{\beta}_j|\bigg] = \sum_{i,j}\langle\widetilde{\beta}_j|\widetilde{\beta}_i\rangle(|\alpha_i\rangle\langle\alpha_j|). \tag{1.46}$$

Confrontano poi (1.45) con (1.46), notiamo che

$$\langle\widetilde{\beta}_i|\widetilde{\beta}_j\rangle = p_i\delta_{ij} \ . \tag{1.47}$$

Risulta allora che la base $|\widetilde{\beta}_i\rangle$ è ortogonale. Possiamo normalizzarla definendo $|\beta_i'\rangle = \dfrac{1}{\sqrt{p_i}}|\widetilde{\beta}_i\rangle$ e questo punto lo stato $|\psi\rangle$ può essere espresso come

$$|\psi\rangle = \sum_i \sqrt{p_i}|\alpha_i\rangle \otimes |\beta_i'\rangle \tag{1.48}$$

L'equazione (1.48) è detta *decomposizione di Schmidt* per uno stato bipartito puro $|\psi\rangle$. Si osservi che è sempre possibile espandere uno stato bipartito puro nella forma esposta, ma tale decomposizione dipende dalla scelta delle basi ortonormali $|\alpha_i\rangle$ e $|\beta_i'\rangle$. Ne consegue che non è invece possibile espandere simultaneamente due stati $|\psi\rangle$ e $|\gamma\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ rispetto alle stesse basi ortonormali di $\mathcal{H}_A$ e $\mathcal{H}_B$.

Usando (1.48) possiamo calcolare ρ^B:

$$\rho^B = Tr_A\big[|\psi\rangle\langle\psi|\big] = \sum_i p_i|\beta_i'\rangle\langle\beta_i'| \ . \tag{1.49}$$

Da ciò si capisce che ρ^A e ρ^B hanno gli stessi *autovalori non nulli*. Se le dimensioni di $\mathcal{H}_A$ e $\mathcal{H}_B$ sono diverse sarà diverso il numero di *autovalori nulli*.

Il numero di autovalori non nulli di ρ^A (o di ρ^B), ovvero il numero di termini non nulli nella decomposizione di Schmidt di $|\psi\rangle$, è detto *numero di Schmidt* e ci da informazioni sul sistema. Un sistema bipartito con numero di Schmidt strettamente maggiore di 1 si

dice *entangled*; in tutti gli altri casi il sistema è detto *separabile*.
Un sistema composto che si trova in uno stato puro e separabile

$$|\psi\rangle = |\phi\rangle_A \otimes |\chi\rangle_B \tag{1.50}$$

da origine agli operatori densità ridotti

$$\rho^A = |\phi\rangle_{AA}\langle\phi| \;\text{ e }\; \rho^B = |\chi\rangle_{BB}\langle\chi| \tag{1.51}$$

che sono anch'essi puri. Uno stato che non è esprimibile come prodotto diretto in questo modo è detto entangled e gli operatori densità ridotti sono stati misti.

Il secondo strumento che è necessario introdurre è la purificazione. Supponiamo di avere un sistema A descritto dall'operatore ρ^A. È sempre possibile introdurre un nuovo sistema R (eventualmente fittizio) e *definire* uno stato puro $|AR\rangle$ per il sistema congiunto AR tale che $\rho^A = Tr_R\big[|AR\rangle\langle AR|\big]$, ovvero lo stato puro $|AR\rangle$ si riduce a ρ^A se ci concentriamo solo sul sistema A. Per dimostrare come tale procedura si possa applicare a qualunque sistema A consideriamo un sistema nello stato ρ^A che ammette una decomposizione spettrale $\rho = \sum_i p_i |i\rangle_{AA}\langle i|$ e consideriamo un sistema R che ha lo stesso spazio degli stati del sistema A e base $|i\rangle_R$. Definiamo uno stato puro del sistema composto AR:

$$|AR\rangle \equiv \sum_i \sqrt{p_i}|i\rangle_A|i\rangle_R \;. \tag{1.52}$$

Calcolando ora l'operatore densità ridotto associato al sistema A troviamo proprio l'operatore ρ^A:

$$\begin{aligned} Tr_R\big[|AR\rangle\langle AR|\big] &= \sum_{i,j} \sqrt{p_i p_j}|i\rangle_{AA}\langle j|Tr\big[|i\rangle_{RR}\langle j|\big] \\ &= \sum_{i,j} \sqrt{p_i p_j}|i\rangle_{AA}\langle j|\delta_{ij} \\ &= \sum_i p_i |i\rangle_{AA}\langle i| = \rho^A. \end{aligned}$$

Capitolo 2

Sistemi quantistici aperti

Lo scopo di questo capitolo è lo studio di un sistema quantistico aperto. In particolare daremo la definizione di superoperatori, usati per descrivere diversi processi che possono avvenire su un sistema quantistico.
Verrà inoltre esposto l'approccio della master equation nella descrizione della dinamica dei sistemi aperti.

2.1 Superoperatori

Il formalismo dei superoperatori, anche noti come operazioni quantistiche, è uno strumento generale per descrivere un'ampia varietà di trasformazioni che un sistema quantistico può subire, ad esempio l'evoluzione temporale unitaria o stocastica, ma anche l'effetto dei processi di misura e di interazione transiente con l'ambiente. La trattazione matematica dei superoperatori è cruciale nella descrizione della dinamica dei *sistemi aperti*.
Come abbiamo esposto nel capitolo precedente uno stato quantistico è descritto da un operatore densità ρ. Un superoperatore è formalmente una trasformazione lineare e positiva che mappa operatori densità in altri operatori densità. In simboli:

$$\rho' = \Gamma(\rho) \tag{2.1}$$

dove ρ è lo stato iniziale prima del processo e $\Gamma(\rho)$ è lo stato finale. Una richiesta ulteriore che viene spesso fatta è che l'operazione quantistica sia *fisica*, ovvero $0 \leq Tr[\Gamma(\rho)] \leq 1 \, \forall \rho$. Qualunque applicazione Γ che soddisfa tali requisiti viene chiamata superoperatore. Due esempi semplici di superoperatori sono le trasformazioni unitarie e le misure, per le quali si ha, rispettivamente, $\Gamma(\rho) = U\rho U^\dagger$ e $\Gamma_m(\rho) = P_m \rho P_m^\dagger$. L'operazione quantistica descrive la dinamica di un processo che avviene sul sistema che inizialmente si trova nello stato ρ.

Grazie ai postulati della meccanica quantistica sappiamo che la dinamica di un sistema chiuso è descritta da una trasformazione unitaria. Un metodo naturale per descrivere

la dinamica di un sistema aperto è quello di considerarlo come generato dall'interazione tra il sistema da studiare, che chiamiamo *sistema principale*, e l'*ambiente*. Insieme essi formano un sistema quantistico chiuso. Il ragionamento che ci porta alla prima definizione di cosa sia un superoperatore è il seguente.
Supponiamo di avere un sistema che si trova inizialmente in uno stato ρ. In generale lo stato finale del sistema, $\Gamma(\rho)$, potrebbe non essere correlato allo stato iniziale tramite una trasformazione unitaria a causa delle interazioni con l'ambiente. Consideriamo allora l'insieme sistema principale-ambiente e supponiamo che il suo stato iniziale sia dato dal prodotto tensoriale $\rho \otimes \rho_E$. Dopo la trasformazione U eseguiamo la traccia parziale sull'ambiente, per ottenere lo stato del sistema principale:

$$\Gamma(\rho) = Tr_E\big[U(\rho \otimes \rho_E)U^\dagger\big]. \tag{2.2}$$

L'evoluzione del sistema principale non risulta più, in generale, unitaria e ne consegue che stati puri possono evolvere in stati misti, un fenomeno che prende il nome di *decoerenza*.

In questa definizione viene fatta un'assunzione importante, ossia che il sistema e l'ambiente inizino in uno stato separabile, anche se in linea generale questo non è vero. I sistemi quantistici interagiscono costantemente con l'ambiente in cui si trovano, stabilendo delle correlazioni ad esempio attraverso lo scambio di calore. Lasciato a se stesso il sistema principale raggiungerà l'equilibrio termico con l'ambiente e ciò instaura delle correlazioni tra i due; tuttavia, in molti casi di interesse pratico, è possibile assumere che il sistema e l'ambiente inizino in uno stato prodotto. Quando gli sperimentali preparano un sistema quantistico in uno stato specifico rimuovono le correlazioni tra il sistema e l'ambiente. Idealmente, la correlazione sarà completamente distrutta, lasciando il sistema in uno stato puro.

Un'altro punto che può destare qualche dubbio è: come può essere definita la trasformazione U dal momento che l'ambiente ha infiniti gradi di libertà? Si scopre che, in realtà, per descrivere una qualsiasi trasformazione $\Gamma(\rho)$ che avviene su un sistema principale di dimensione d basta considerare un che lo spazio di Hilbert dell'ambiente abbia dimensione non maggiore di d^2[4].

2.2 Rappresentazione di Kraus

I superoperatori possono essere descritti in una elegante forma conosciuta come *rappresentazione di Kraus*, o *operator-sum representation*, che è essenzialmente una riscrittura dell'equazione (2.2) in termini espliciti degli operatori che agiscono solo sull'operatore densità del sistema principale. Sia $|e_k\rangle$ una base ortonormale per lo spazio degli stati, che consideriamo di dimensione finita, dell'ambiente. Sia $\rho_E = |e_0\rangle\langle e_0|$ il suo stato iniziale. Non è restrittivo assumere che l'ambiente inizi in uno stato puro, dal momento che è sempre possibile introdurre un altro sistema per purificarlo. L'equazione (2.2) può così

essere riscritta come:

$$\begin{aligned}\Gamma(\rho) &= \sum_k \langle e_k|U\big[\rho\otimes|e_0\rangle\langle e_0|\big]U^\dagger|e_k\rangle \\ &= \sum_k E_k\rho E_k^\dagger \end{aligned} \tag{2.3}$$

dove $E_k \equiv \langle e_k|U|e_0\rangle$ sono detti *operatori di Kraus* e agiscono sullo spazio degli stati del sistema principale mentre l'equazione (2.3) è nota con il nome di *rappresentazione di Kraus*. Il *teorema di Choi-Kraus* assicura che per ogni operazione quantistica Γ esiste una sua rappresentazione nella forma (2.3), detta *rappresentazione di Kraus*.
Gli operatori di Kraus $\{E_k\}$soddisfano un'importante costrizione conosciuta come *relazione di completezza*. Nel caso classico, la relazione di completezza nacque dall'esigenza di avere probabilità che si sommano all'unità. Nel caso quantistico la relazione di completezza diventa

$$\begin{aligned} 1 &= Tr\big[\Gamma(\rho)\big] \\ &= Tr\bigg[\sum_k E_k\rho E_k^\dagger\bigg] \\ &= Tr\bigg[\sum_k E_k^\dagger E_k\rho\bigg]. \end{aligned} \tag{2.4}$$

Dal momento che questa relazione è vera per ogni ρ, segue che dobbiamo avere

$$\sum_k E_k^\dagger E_k = I. \tag{2.5}$$

Questa equazione è soddisfatta dai soli superoperatori che preservano la traccia. Come la scelta della base $|e_k\rangle$, anche la rappresentazione di Kraus è arbitraria e diverse rappresentazioni sono legate da una trasformazione unitaria. Siano $\{M_i\}$ e $\{N_j\}$ due rappresentazioni dello stesso superoperatore relative alle basi $|m_i\rangle$ e $|n_j\rangle$ dello spazio degli stati dell'ambiente; si ha allora $M_i = U_{ij}N_j$, dove U_{ij} è la matrice unitaria che che soddisfa $|m_i\rangle = \sum_{i,j} U_{ij}|n_j\rangle$.
Esistono anche superoperatori che non preservano la traccia, per le quali $\sum_k E_k^\dagger E_k \leq I$ che descrivono processi in cui informazioni aggiuntive su quello che si è verificato nel processo sono ottenute tramite misurazioni. La mappa Γ nella forma (2.3) per cui $\sum_k E_k^\dagger E_k \leq I$ fornisce una definizione alternativa di un superoperatore.

Per dare una interpretazione fisica a questo formalismo consideriamo la situazione in cui si effettua una misura sull'ambiente nella base $|e_k\rangle$, dopo l'azione della trasformazione unitaria U sul sistema composto che, inizialmente si trova nello stato $\rho\otimes|e_0\rangle\langle e_0|$. Lo stato del sistema principale dopo la misura, nel caso in cui il risultato sia k, è:

$$\begin{aligned}\rho_k &\propto Tr_E\big[|e_k\rangle\langle e_k|U(\rho\otimes|e_0\rangle\langle e_0|)U^\dagger|e_k\rangle\langle e_k|\big] \\ &= \langle e_k|U(\rho\otimes|e_0\rangle\langle e_0|)U^\dagger|e_k\rangle = E_k\rho E_k^\dagger \end{aligned} \tag{2.6}$$

normalizzandolo otteniamo

$$\rho_k = \frac{E_k \rho E_k^\dagger}{Tr[E_k \rho E_k^\dagger]} \ . \tag{2.7}$$

La probabilità di ottenere k come risultato è allora $p(k) = Tr[E_k \rho E_k^\dagger]$, perciò si ha:

$$\Gamma(\rho) = \sum_k E_k \rho E_k^\dagger = \sum_k p(k) \rho_k \ . \tag{2.8}$$

Tramite questo ragionamento possiamo concludere che l'azione del superoperatore definito dagli operatori di Kraus $\{E_k\}$ è quella di sostituire a ρ uno dei ρ_k con probabilità data da $p(k)$.

Estendiamo ora questo formalismo al caso più generale in cui non si ha conservazione della traccia. Consideriamo la situazione in cui il sistema principale Q si trova nello stato ρ e l'ambiente E si trova nello stato σ. Lo stato del sistema composto è $\rho^{QE} = \rho \otimes \sigma$. Dopo l'evoluzione unitaria U effettuiamo una misura descritta dai proiettori P_n. La situazione in cui non venga fatta alcuna misura è quella in cui il solo risultato della misura è $n = 0$ ed è descritta dall'operatore $P_0 \equiv I$. Lo stato finale di QE è

$$\frac{P_n U(\rho \otimes \sigma) U^\dagger P_n}{Tr[P_n U(\rho \otimes \sigma) U^\dagger P_n]} \tag{2.9}$$

nel caso in cui il risultato della misura sia n. Eseguendo la traccia parziale su E otteniamo lo stato finale di Q. Definendo $\Gamma_n(\rho) \equiv Tr_E[P_n U(\rho \otimes \sigma) U^\dagger P_n]$ lo stato finale di Q risulta essere

$$\frac{\Gamma_n(\rho)}{Tr[\Gamma_n(\rho)]} \tag{2.10}$$

dove $Tr[\Gamma_n(\rho)]$ risulta la probabilità di avere n come risultato.
Sia $\sigma = \sum_j q_j |j\rangle\langle j|$ la decomposizione dello stato di E. Introducendo una base $|e_k\rangle$ di E abbiamo:

$$\begin{aligned} \Gamma_n(\rho) &= \sum_{j,k} q_j Tr_E \left[|e_k\rangle\langle e_k| P_n U(\rho \otimes |j\rangle\langle j|) U^\dagger P_n |e_k\rangle\langle e_k| \right] \\ &= \sum_{j,k} E_{jk} \rho E_{jk}^\dagger \end{aligned} \tag{2.11}$$

dove $E_{jk} \equiv \sqrt{q_j} \langle e_k | P_n U | j \rangle$. Questa formula fornisce un metodo esplicito per il calcolo della rappresentazione di Kraus di un superoperatore Γ_n una volta che sono noti lo stato iniziale σ di E e la trasformazione U che regola la dinamica tra il sistema principale e l'ambiente.

Possiamo mostrare come l'applicazione traccia $\rho \to Tr[\rho]$ sia effettivamente un superoperatore. Consideriamo uno spazio di Hilbert $\mathcal{H}_Q$ generato dalla base ortonormale

$|1\rangle ... |d\rangle$, e sia $\mathcal{H}'_Q$ uno spazio unidimensionale generato dal vettore $|0\rangle$. Definiamo

$$\Gamma(\rho) \equiv \sum_{i=1}^{d} |0\rangle\langle i|\rho|i\rangle\langle 0|. \tag{2.12}$$

Γ definisce una operazione quantistica in quanto ammette una rappresentazione di Kraus. Si noti che $\Gamma(\rho) = Tr[\rho]|0\rangle\langle 0|$, per cui, a meno del fattore moltiplicativo ininfluente $|0\rangle\langle 0|$ tale superoperatore è identico alla applicazione traccia. Possiamo spingerci oltre osservando che anche la traccia parziale è una operazione quantistica. Sia QR un sistema congiunto e immaginiamo di voler farne la traccia parziale sul sistema R. Sia $|j\rangle$ una base di per il sistema R. Definiamo allora $E_i : \mathcal{H}_Q R \to \mathcal{H}_Q$ tramite la relazione:

$$E_i\left(\sum_j \lambda_j |q_j\rangle\langle j|\right) \equiv \lambda_i |q_i\rangle \ , \tag{2.13}$$

dove λ_j sono numeri complessi e $|q_j\rangle$ sono stati arbitrari di Q. Scegliamo gli $\{E_i\}$ come operatori di Kraus del superoperatore Γ:

$$\Gamma(\rho) = \sum_i E_i \rho E_i^\dagger \ . \tag{2.14}$$

Abbiamo definito in questo modo una operazione quantistica dal sistema QR al sistema Q. Notiamo ora che

$$\Gamma(\rho \otimes |j\rangle\langle j'|) = \rho\delta_{ji'} = Tr_R\big[\rho \otimes |j\rangle\langle j'|\big] \ , \tag{2.15}$$

dove ρ è un operatore Hermitiano sullo spazio Q, $|j\rangle$ e $|j'\rangle$ sono elementi della base di R. Per la linearità di Γ e Tr_R si ha $\Gamma = Tr_R$.

2.3 Master equation

Il formalismo dei superoperatori ci fornisce un metodo per descrivere l'evoluzione delle matrici densità nello stesso modo in cui le trasformazioni unitarie forniscono una descrizione dell'evoluzione dei sistemi chiusi. In quest'ultimo caso conviene caratterizzare il sistema con un un Hamiltoniano che descrive l'evoluzione temporale su intervalli di tempo infinitesimi. La dinamica viene descritta da una equazione differenziale, l'equazione di Schrödinger, che possiamo integrare per calcolare l'evoluzione su intervalli di tempo finiti.

Allo stesso modo è spesso possibile descrivere l'evoluzione, non necessariamente coerente, degli operatori densità tramite una equazione differenziale, perlomeno in buona approssimazione. Tale descrizione è possibile solo nel caso in cui l'evoluzione temporale

del sistema risulta *Markoviana*. Un processo Markoviano è un processo in cui la probabilità che il sistema passi in un determinato stato dipende *solo* dallo stato del sistema all'istante immediatamente precedente e non da *come* si è giunti a questo stato. Infatti, se l'evoluzione di $\rho(t)$ è governata da una equazione differenziale al primo ordine in t, allora $\rho(t+dt)$ è completamente determinato da $\rho(t)$.

Abbiamo visto che è sempre possibile descrivere l'evoluzione dell'operatore densità ρ^A nello spazio di Hilbert $\mathcal{H}_A$ immaginandola come una evoluzione unitaria nello spazio di Hilbert esteso $\mathcal{H}_A \otimes \mathcal{H}_E$, dove $\mathcal{H}_E$ rappresenta lo spazio di Hilbert relativo all'ambiente. Il fatto che l'evoluzione in $\mathcal{H}_A \otimes \mathcal{H}_E$ possa essere descritta tramite l'equazione di Schrödinger non ci garantisce che l'evoluzione di $\rho^A(t)$ sia Markoviana. Il punto cruciale è che la conoscenza di $\rho^A(t)$ non basta come condizione iniziale per la nostra equazione differenziale: è necessario conoscere anche lo stato iniziale dell'ambiente. In questa situazione $\rho^A(t+dt)$ dipende da $\rho^A(t)$, ma anche da ρ^A in istanti di tempo precedenti dal momento che l'ambiente conserva temporaneamente *memoria* dell'informazione ricevuta e può trasferirla indietro al sistema A. Un sistema aperto è quindi dissipativo perché l'informazione può passare dal sistema di interesse all'ambiente e viceversa, producendo fluttuazioni non-Markoviane nel sistema. Ad eccezione del caso di evoluzione unitaria, queste fluttuazioni sono inevitabili e non è possibile fare una descrizione esattamente Markoviana della meccanica del nostro sistema quantistico. In molti contesti la descrizione Markoviana risulta una buona approssimazione. Il punto chiave è che è necessario avere una separazione netta tra il tempo tipico di correlazione e la scala temporale dell'evoluzione che vogliamo seguire. Se indichiamo con $(\Delta t)_{res}$ il tempo che l'ambiente impiega per disperdere completamente l'informazione ricevuta dal sistema, allora siamo sicuri che dopo un tempo $(\Delta t)_{coarse}$ siamo sicuri che l'informazione non può tornare indietro dal sistema all'ambiente per influenzare la successiva evoluzione del sistema. Per descrivere l'evoluzione del nostro sistema studiamo la dinamica ad intervalli di tempo prestabiliti $(\Delta t)_{coarse}$. Sostanzialmente è come guardare il sistema attraverso un filtro che taglia le frequenze $\omega \gg (\Delta t)^{-1}_{coarse}$. Una descrizione approssimativamente Markoviana del sistema è possibile, in quest'ottica, solo se $(\Delta t)_{res} \ll (\Delta t)_{coarse}$; ovvero se possiamo trascurare la memoria dell'ambiente. In pratica questa condizione è spesso verificata. Si pensi ad esempio alla fisica atomica in cui sia ha $(\Delta t)_{res} \sim \hbar/k_B T \sim 10^{-14} s$ (dove k_B è la costante di Boltzmann e T indica la temperatura) che è diversi ordini di grandezza più grande del tempo di vita medio di uno stato atomico eccitato.

Come abbiamo già visto l'evoluzione unitaria dell'operatore densità è descritta dall'equazione di Schrödinger (1.23) che ammette come soluzione

$$\rho(t) = e^{-iHt}\rho(0)e^{iHt} \tag{2.16}$$

se H non dipende dal tempo. Il nostro obiettivo è quello di generalizzare questa equazione al caso in cui l'evoluzione è Markoviana, ma non unitaria per cui si ha

$$\dot{\rho} = \mathcal{L}[\rho(0)]. \tag{2.17}$$

L'operatore lineare $\mathcal{L}$, che genera un superoperatore nello stesso senso in cui H genera l'evoluzione temporale unitaria, è detto *Lindbladiano*. Se tale operatore non dipende dal tempo una soluzione formale di tale equazione è $\rho(t) = e^{\mathcal{L}t}[\rho(0)]$.
Sfruttando l'approssimazione Markoviana è possibile ricavare la *Master Equation* nella sua *forma Lindbladiana*[5]:

$$\dot{\rho} = -\frac{i}{\hbar}[H, \rho] + \sum_j \left[2L_j \rho L_j^\dagger - \left\{ L_j^\dagger L_j, \rho \right\} \right], \tag{2.18}$$

dove $\{X, Y\} = XY + YX$ denota l'anticommutatore di due operatori X e Y, H è l'Hamiltoniano del sistema e gli L_j sono gli *operatori di Lindblad* che rappresentano l'interazione tra il sistema e l'ambiente.

L'approccio della Master Equation è meno generale del formalismo dei superoperatori. La soluzione della Master Equation determina la dipendenza temporale di ρ. Una volta nota la funzione $\rho(t)$ è possibile scrivere la sua rappresentazione di Kraus:

$$\Gamma_t(\rho(0)) = \rho(t) = \sum_k E_k(t) \rho(0) E_k^\dagger(t), \tag{2.19}$$

con $\Gamma_{t=0} = I$ e dove gli $E_k(t)$ sono operatori di Kraus dipendenti dal tempo determinati dalla soluzione dell'equazione. Invece non è detto che un processo quantistico con una certa rappresentazione di Kraus possa essere scritto come una equazione differenziale. Infatti isuperoperatori possono descrivere anche processi non-Markoviani dal momento che descrivono solo i cambiamenti di stato del sistema e non la sua evoluzione temporale continua.

Capitolo 3

Entropia e informazione

In queste ultime pagine, facendo uso degli strumenti matematici esposti nei capitoli precedenti, viene introdotto il concetto di entropia, esplicitandone il legame con la termodinamica, per poi approcciare la teoria dell'informazione classica e quantistica.

3.1 Entropia di Shannon

La teoria classica dell'informazione studia con un approccio matematico la misurazione e la trasmissione di informazione attraverso un canale di comunicazione. Lo strumento chiave per descrivere l'informazione classica è l'*entropia di Shannon*. Supponiamo di apprendere il valore di una certa variabile casuale X. L'entropia di Shannon ci dice quanta informazione acquisiamo, in media, *dopo* aver appreso il valore di X. Una interpretazione alternativa è la seguente: l'entropia di X misura l'*incertezza* riguardo a X *prima* che di saperne il valore. L'entropia dipende solo dalle probabilità associate ai possibili valori di una certa variabile, ma non dipende dai valori stessi che essa può assumere. L'entropia di Shanno associata a X viene spesso espressa come funzione delle probabilità dei possibili esiti $p_1, ..., p_n$ ed è definita da

$$H(X) \equiv H(p_1, ..., p_n) = -\sum_x p_x \log(p_x) \tag{3.1}$$

dove usiamo 'log' per indicare il logaritmo in base due nel resto di questa trattazione. Notiamo che il caso in cui $p_x = 0$ potrebbe creare problemi, ma eventi impossibili non possono contribuire all'entropia per cui assumiamo $0 \log 0 \equiv 0$. D'altronde il fatto che $\lim_{x\to 0} x \log x$ da' supporto alla nostra convenzione. Questa formula è giustificata dal fatto che può essere usata per quantificare le risorse necessarie per immagazzinare l'informazione. In particolare, se consideriamo una sorgente che emette stringhe $X_1, X_2, ...$ di variabili casuali indipendenti e identicamente distribuite. L'entropia $H(X) \equiv H(X_1) = H(X_2) = ...$ rappresenta il numero minimo di bit necessari per un

simbolo dalla sorgente. Questo risultato è conosciuto come primo teorema di Shannon (o *Shannon's noiseless coding theorem*).

Per chiarire le idee consideriamo un esempio. Supponiamo che una sorgente di informazione produca uno dei quattro simboli $1, 2, 3$ o 4. Per immagazzinarli, in generale, servono due bit ogni volta che la sorgente produce un segnale. Se, però, il simbolo 1 è prodotto con probabilità $1/2$, il simbolo 2 con probabilità $1/4$ e i simboli 3 e 4 entrambi con probabilità $1/8$ allora possiamo ridurre il numero di bit necessari, in linea di principio, per memorizzare il segnale emesso associando un codice ad ciascun simbolo. Un possibile modo per comprimere è quello di codificare 1 come la stringa di bit 0, 2 tramite la stringa 10, 3 come 110 e 4 come 111. In questo modo la lunghezza media di una stringa è di $7/4$ bit che è minore dei due bit richiesti in assenza di codifica. Calcolando l'entropia associa alla sorgente tramite la formula (3.1) troviamo lo stesso valore che si dimostra essere anche il valore massimo di compressione che può teoricamente essere raggiunto.

L'entropia di una variabile che può assumere due soli valori è detta *binaria*:

$$H_{bin}(p) \equiv -p\log(p) - (1-p)\log(1-p) = H_{bin}(1-p) \tag{3.2}$$

dove p e $1-p$ sono le probabilità dei due possibili esiti. Il massimo di questa funzione di ha per $p = 1/2$.

La funzione H ha la proprietà di essere una funzione *concava*:

$$H(\lambda p_1 + (1-\lambda)p_2) \geq \lambda H(p_1) + (1-\lambda)H(p_2) \tag{3.3}$$

dove $0 \leq \lambda \leq 1$

Un altro strumento utile è l'*entropia relativa* che misura la *vicinanza* di due distribuzioni di probabilità $p(x)$ e $q(x)$ sullo stesso insieme di variabili x. L'entropia relativa di p e q, distribuzioni di probabilità sullo stesso insieme di variabili x è

$$H(p(x)||q(x)) = \sum_x p(x)\log\,\frac{p(x)}{q(x)} \equiv -H(X) - p(x)\log\,q(x)\ . \tag{3.4}$$

La motivazione che ci permette di interpretare l'entropia relativa come una misura della vicinanza di due distribuzioni è data da un teorema che afferma che $H(p(x)||q(x)) \geq 0$ ccon l'uguaglianza che si verifica se e solo se $p(x) = q(x)\ \forall x$. Una conseguenza di questo teorema è che l'entropia di un set di variabili uniformemente distribuite è massima. Consideriamo ad esempio una distribuzione di probabilità $p(x)$ per l'insieme X con d possibili esiti diversi e sia $q(x) \equiv 1/d$ la distribuzione di probabilità uniforme. Si ha

$$H(p(x)||q(x)) = H(p(x)||1/d) = \log\,d - H(X). \tag{3.5}$$

Dalla non negatività dell'entropia relativa otteniamo $H(X) \leq \log\,d$.

Consideriamo ora due variabili casuali X e Y. Definiamo l'*entropia congiunta* di X e Y come

$$H(X,Y) \equiv -\sum_{x,y} p(x,y)\log\,p(x,y) \tag{3.6}$$

Questa quantità misura la nostra incertezza riguardo la coppia (X, Y). Se ora immaginiamo di aver appreso il valore di Y, ovvero abbiamo acquisito $H(Y)$ bit di informazione, l'incertezza che rimane sulla coppia (X, Y) è detta *entropia condizionale* di X una volta noto Y e vale

$$H(X|Y) \equiv H(X, Y) - H(Y). \tag{3.7}$$

È possibile mostrare che vale che $H(X, Y) \leq H(X) + H(Y)$ con l'uguaglianza che si verifica nel caso in cui X e Y siano variabili casuali indipendenti. Tale proprietà è detta *subadditività* ed è dovuta al fatto che parte dell'informazione in XY è contenuta nelle correlazioni tra X e Y. Vale anche una proprietà detta *subadditività forte*: $H(X, Y, Z) + H(Y) \leq H(X, Y) + H(Y, Z)$ dove l'uguaglianza è valida se e solo se $Z \to Y \to X$ forma una catena di Markov, ovvero una sequenza di variabili casuali $X_1 \to X_2 \to ...$ tale che X_{n+1} è indipendente da $X_1, .., X_{n-1}$, dato X_n. Introduciamo anche un'altra quantità, detta *informazione mutua* di X e Y, che descrive quanta informazione X e Y hanno in comune. Se alla somma delle entropie di X e Y sottraiamo l'informazione comune, ovvero l'entropia congiunta, di (X, Y) otteniamo l'informazione mutua:

$$H(X : Y) \equiv H(X) + H(Y) - H(X, Y) = H(X) - H(X|Y). \tag{3.8}$$

Le relazioni tra i diversi tipi di entropie sono le seguenti:

1) $H(X, Y) = H(Y, X),\ H(X : Y) = H(Y : X)$;

2) $H(Y|X) \geq 0$ da cui si ricava anche $H(X : Y) \leq H(Y)$ con l'uguaglianza valida se e solo se $Y = f(X)$ con f funzione arbitraria;

3) $H(X) \leq H(X, Y)$ con l'uguaglianza valida se e solo se $Y = f(X)$;

4) $H(Y|X) \leq H(Y)$ da cui $H(X : Y) \geq 0$ con l'uguaglianza valida se e solo se X e Y sono variabili casuali indipendenti.

v) forse p536 nielsen chaining rule for conditional entropies.

3.2 Entropia di von Neumann

Come abbiamo visto nella precedente sezione, l'entropia di Shannon quantifica l'incertezza associata ad una distribuzione di probabilità classica. Gli stati quantistici sono descritti in modo simile sostituendo alle distribuzione di probabilità le matrici densità. Definiamo l'*entropia di von Neumann* come segue

$$S(\rho) \equiv -Tr[\rho \log\rho], \tag{3.9}$$

dove abbiamo sempre usato il logaritmo in base due. Se λ_x sono gli autovalori di ρ allora vale

$$S(\rho) = -\sum_x \lambda_x \log(\lambda_x), \tag{3.10}$$

dove valgono le stesse assunzioni fatte in precedenza nel caso in cui $\lambda_x = 0$. Consideriamo un semplice esempio: l'operatore densità di uno spazio d-dimensionale in cui i diversi stati sono equiprobabili è $\rho = I/d$ e la sua entropia è $\log(d)$. Come nel caso classico si possono definire l'entropia relativa quantistica $S(\rho||\sigma) = S(\rho) - S(\sigma)$, per cui vale la *disuguaglianza di Klein*: $S(\rho||\sigma) \geq 0$ con l'uguaglianza valida se e solo se $\rho = \sigma$. Le proprietà dell'entropia di von Neumann sono le seguenti:

1) Per uno stato è puro $\rho = |\psi\rangle\langle\psi|$ si ha $S(\rho) = 0$.

2) Se U è una matrice unitaria allora vale $S(\rho) = S(U\rho U^{-1})$.

3) In uno spazio d-dimensionale $S(\rho) \leq \log d$ e l'uguaglianza vale se e solo se lo stato è massimamente mescolato, ovvero quando $\rho = I/d$.

4) $\forall\ \lambda_1, \lambda_2, ..., \lambda_n \geq 0$ tali che $\lambda_1 + \lambda_2 + ... + \lambda_n = 1$ si ha

$$S(\lambda_1\rho_1 + ... + \lambda_n\rho_n) \geq \lambda_1 S(\rho_1) + ... + \lambda_n S(\rho_n). \tag{3.11}$$

Ovvero l'entropia di von Neumann è una funzione concava.

5) Supponiamo di misurare una osservabile $A = \sum_x |a_x\rangle a_x \langle a_x|$ di uno stato ρ. La probabilità di misurare a_x è $p(a_x) = \langle a_x|\rho|a_x\rangle$. Calcolando l'entropia di Shannon per l'ensemble $X = \{a_x, p(a_x)\}$ si ha

$$H(X) \geq S(\rho). \tag{3.12}$$

L'uguaglianza vale se e solo se l'osservabile A commuta con ρ. Come conseguenza, scegliere una *buona* osservabile consente di fare previsioni statistiche più accurate.

6) Sia ρ la matrice densità associata all'ensemble $X = \{|\phi_x\rangle, p_x\}$.

$$H(X) \geq S(\rho), \tag{3.13}$$

con l'uguaglianza che si verifica se e solo se gli stati $|\phi_x\rangle$ sono ortogonali. Fisicamente significa che gli stati non ortogonali sono *indistinguibili* una volta mescolati.

7) Per un sistema bipartito AB vale

$$S(\rho^{AB}) \leq S(\rho^A) + S(\rho^B); \tag{3.14}$$

dove le matrici densità ridotte si calcolano come di consueto tramite la traccia parziale. L'uguaglianza si ha solo nel caso in cui $\rho^{AB} = \rho^A \otimes \rho^B$. Questa proprietà è detta *subadditività* e ci dice sostanzialmente che l'entropia è *additiva* per sistemi non correlati.

8) Siano A, B e C tre sistemi quantistici. Vale la seguente:

$$S(A,B,C) + S(B) \leq S(A,B) + S(B,C). \tag{3.15}$$

Questa proprietà è nota con il nome di *subadditività forte*. Possiamo notare che valgono anche le seguenti:

i) Scartare un sistema quantistico non può aumentare l'informazione mutua:

$$S(A:B) \leq S(A:B,C); \tag{3.16}$$

ii) I superoperatori non aumentano l'informazione mutua. Sia Γ un superoperatore che preserva la traccia. Denotando con A' e B' i sistemi A e B dopo l'azione di Γ si ha, in fomrule:

$$S(A':B') \leq S(A:B). \tag{3.17}$$

9) Per un sistema bipartito vale la disuguaglianza triangolare:

$$S(\rho^{AB}) \geq |S(\rho^A) - S(\rho^B)|. \tag{3.18}$$

È opportuno notare che questa proprietà dell'entropia di von Neumann è in contrasto con la proprietà 3) dell'entropia di Shannon. Intuitivamente questa proprietà dell'entropia classica ci dice che c'è più informazione nel sistema composto che nelle parti che lo compongono. Ciò non cale per l'entropia di von Neumann. Basta considerare, ad esempio, lo stato di Bell (1.40) di due qubit AB. Come abbiamo già discusso tale stato è puro e dunque $S(A,B) = 0$. D'altra parte si ha $S(\rho^A = I/2) = 1$. Alternativamente possiamo affermare che la quantità $S(B|A)$ è negativa.

10) Sia $\rho = \sum_i p_i \rho_i$ Con p_i probabilità e ρ_i operatori densità. Vale allora:

$$S(\rho) \leq \sum_i p_i S(\rho_i) + H(p_i), \tag{3.19}$$

dove l'uguaglianza vale solo se i ρ_i hanno supporto su sottospazi ortogonali.

Come nel caso classico abbiamo definito l'*informazione mutua*:

$$\begin{aligned} S(A:B) &\equiv S(A) + S(B) - S(A,B) \\ &= S(B) - S(B|A) = S(A) - S(A|B). \end{aligned} \tag{3.20}$$

Dove abbiamo definito $S(A,B) \equiv -Tr\left[\rho^{AB} \log \rho^{AB}\right]$ dove ρ^{AB} rappresenta la matrice densità del sistema AB. Altre proprietà possono essere trovate nella referenza [5].

3.3 Entropia e termodinamica

Il concetto di entropia fece la sua prima comparsa nello studio della termodinamica. È interessante notare come le proprietà di $S(\rho)$ abbiano importanti implicazioni termodinamiche.

Consideriamo un sistema A e l'ambiente in cui questo si trova E. Se inizialmente sono non correlati si troveranno nello stato $\rho^{AE} = \rho^A \otimes \rho^E$ e l'entropia risulta additiva per la proprietà 7), ovvero $S(\rho^{AE}) = S(\rho^A) + S(\rho^E)$. Il sistema, che è chiuso, evolverà seguendo una trasformazione unitaria U_{AE}:

$$\rho^{AE} \to \rho^{AE'} = U_{AE}\rho^{AE}U_{AE}^{-1}, \tag{3.21}$$

e per la proprietà 2) si ha $S(\rho^{AE'}) = S(\rho^{AE})$ e usando la subadditività sullo stato $\rho^{AE'}$ troviamo

$$S(\rho^A) + S(\rho^E) = S(\rho^{AE'}) \leq S(\rho^{A'}) + S(\rho^{E'}). \tag{3.22}$$

Definendo l'entropia dell'universo come somma delle entropie del sistema e dell'ambiente possiamo concludere che *l'entropia dell'universo non può mai diminuire*. Questa è una possibile formulazione del secondo principio della termodinamica che abbiamo ricavato assumendo che inizialmente il sistema e l'ambiente fossero non correlati. In generale il sistema e l'ambiente interagiscono e ciò fa si che l'entropia totale sia effettivamente *crescente*.

Da un punto di vista microscopico, la nostra conoscenza iniziale dello stato del sistema, ovvero la possibilità di distinguere diversi stati iniziali, in principio ortogonali, è perduta e viene 'codificata' nell'entanglement tra il sistema e l'ambiente. Teoricamente questa informazione può essere totalmente recuperata, ma in pratica risulta inaccessibile a osservatori spazialmente locali. Ciò risulta nella *non reversibilità* termodinamica.

Nella discussione sulla Master Equation abbiamo detto che tipicamente l'ambiente 'dimentica' velocemente l'informazione ricevuta dal sistema e per questo motivo, se la nostra discretizzazione del tempo è sufficientemente grezza, possiamo considerare il sistema e l'ambiente come non correlati a qualunque istante di tempo (da cui l'approssimazione Markoviana). Sotto queste ipotesi l'entropia cresce monotonicamente rispettando le leggi fisiche di conservazione.

Nel 1871, James Clerk Maxwell propose un ipotetico dispositivo in grado di far diminuire l'entropia di un sistema. Lui immaginava un essere in miniatura, il '*diavoletto di Maxwell*', che fosse in grado di diminuire l'entropia di un gas in un contenitore chiuso, inizialmente all'equilibrio, separando individualmente le molecole veloci da quelle lente nelle due metà separate del contenitore. Il diavoletto si trova vicino ad una porta scorrevole che può aprire a piacimento. Quando una molecola veloce si avvicina dal lato sinistro il diavoletto apre la porta e la fa passare dall'altro lato per poi richiudere la porta. Ripetendo questo procedimento più volte si può teoricamente diminuire l'entropia del sistema, contraddicendo apparentemente il secondo principio della termodinamica.

La soluzione a questo paradosso è fornita dal Principio di Landauer di cui parliamo nella prossima sezione.

3.4 Energia e computazione

Nella moderna teoria della computazione i problemi principali riguardano le il tempo e lo spazio necessario per risolvere un determinato problema (*teoria della complessità computazionale*). Un'altra risorsa importante per la computazione è l'*energia* richiesta per la risoluzione di un certo problema. Sorprendentemente di scopre che la computazione, classica o quantistica, può essere svolta senza consumo di energia e ciò è strettamente correlato con la *reversibilità* della computazione. Se consideriamo le porte logiche possiamo osservare che spesso non sono reversibili. Si pensi alla porta NAND che ha due bit come input e uno come output. Lo stato in ingresso non può essere univocamente determinato dalla conoscenza del bit in uscita. Infatti si ha:

$$\begin{aligned} 0 \text{ NAND } 0 &= 1 \\ 0 \text{ NAND } 1 &= 1 \\ 1 \text{ NAND } 0 &= 1 \\ 1 \text{ NAND } 1 &= 0, \end{aligned} \tag{3.23}$$

e allo stato di output 1 possono corrispondere diversi stati di input. La porta NAND è dunque *irreversibile*. La porta NOT, al contrario, è reversibile.

Un modo per intendere la reversibilità computazionale è intendendola in termini di *perdita di informazione*. Se una porta logica è irreversibile allora deve essere stata necessariamente eliminata dell'informazione in ingresso. Contrariamente, nelle porte logiche irreversibili, l'informazione non viene perduta dal momento che lo stato in input può sempre essere recuperato tramite lo stato di output. Il collegamento tra la reversibilità della computazione e la perdita di informazione è dato dal **Principio di Landauer**. Esistono due forme di questo principio che sono analoghe.

Prima forma: Supponiamo che un computer elimini un bit di informazione. L'energia dissipata nell'ambiente è *almeno* pari a $k_B T \ln 2$, dove k_B è la costante di Boltzmann e T è la temperatura dell'ambiente in cui si trova il computer.

Seconda forma: Supponiamo che un computer elimini un bit di informazione. L'entropia dell'ambiente cresce *almeno* di $k_B \ln 2$, dove k_B è la costante di Boltzmann.

Al giorno d'oggi i computer esistenti dissipano più energia del minimo teorico previsto da Landauer. Questo principio illumina un punto di vista interessante: se tutta la

computazione potesse essere svolta in modo reversibile allora, teoricamente, i computer potrebbero senza dissipare energia, dal momento che nessun bit viene cancellato. Ma è effettivamente possibile fare una computazione universale senza cancellare informazione? La risposta *deve* essere sì, dal momento che le leggi fisiche sono fondamentalmente reversibili.

Un modo per capire intuitivamente il Principio di Landauer è tramite il seguente esperimento mentale. Supponiamo di avere una molecola in un contenitore. Posso interpretare tale molecola come informazione immaginando che rappresenti un bit che assume valore 0 se la molecola si trova nel lato sinistro del contenitore e 1 se si trova nel lato destro. Nonostante si tratti di un gas con una sola molecola possiamo applicare comunque le leggi della termodinamica. Immaginando di comprimere il gas fino a dimezzarne il volume possiamo calcolare la variazione di entropia relativa a questo processo. Usando le equazioni classiche della termodinamica

$$\begin{aligned} dE &= \delta Q + \delta L \\ \delta S &= \frac{\delta Q}{T}, \end{aligned} \tag{3.24}$$

dove T rappresenta la temperatura del bagno termico in cui si trova il contenitore, possiamo calcolare il lavoro svolto durante questa trasformazione isoterma integrando $\delta L = -pdV$ tra V e $V/2$, dove V rappresenta il volume del contenitore e p la pressione, intesa in questo contesto come la media temporale su molteplici urti della molecola contro le pareti. Troviamo quindi $L = k_B T \ln 2$ da cui si ha $\Delta S_{gas} = k_B T \ln 2$. Dal momento che l'entropia totale non può decrescere si ha $\Delta S_{amb} \geq k_B T \ln 2$, in accordo con il Principio[1].

Come anticipato questo Principio permette di risolvere il paradosso del diavoletto di Maxwell. Per dividere le molecole veloci da quelle lente, il diavoletto deve effettuare misure per determinare la velocità delle molecole. Il risultato di tale misura deve essere in qualche modo immagazzinato nella memoria del diavoletto. Dal momento che nessuna memoria è illimitata il demone deve, prima o poi, eliminare informazione per poterne immagazzinare di nuova. Per il Principio di Landauer l'atto di cancellazione di informazione produce un aumento di entropia che è almeno pari alla diminuzione di entropia causata dall'azione del diavoletto assicurando così la validità del secondo principio della termodinamica.

Il Principio può essere letto in una ottica diversa: una operazione logica reversibile può, teoricamente, essere eseguita con un sistema fisico che lavora in un modo termodinamicamente reversibile. Genericamente, però, anche una operazione logica irreversibile può essere interpretata come un processo termodinamico reversibile poiché rappresenta il trasferimento, reversibile, di entropia dal sistema all'ambiente. A questo proposito contro il Principio di Landauer sono state mosse numerose obiezioni, principalmente a causa del legame che tale principio ha con la seconda legge della termodinamica. In particolare, dal momento che il principio non risulta indipendente dalla seconda legge, ciò significa che è insufficiente o non necessario per risolvere il paradosso di Maxwell.

Altre critiche affermano che il principio è essenzialmente falso, o privo di significato, e si basano principalmente sui seguente punti:

1) Dal momento che non esiste un legame tra le grandezze termodinamiche, come calore e lavoro, e la reversibilità delle operazioni logiche il principio è privo di significato;

2) È falso dal momento che *tutte* le manipolazioni di dati, reversibili o non, richiedono una dissipazione di energia almeno pari a $k_B T \ln 2$, e in genere anche molto di più, per essere portata a termine da un apparato fisico;

3) È falso perché, in linea teorica, è possibile svolgere operazioni logiche irreversibili in un modo termodinamicamente reversibile.

Nonostante queste critiche l'importanza del principio resta indiscussa. Permette infatti di salvare in modo esattamente sufficiente il secondo principio della termodinamica. Il Principio ha permesso di spiegare *perché* il diavoletto non può funzionare senza affidarsi alla semplice spiegazione "non può funzionare perché viola il secondo principio". Per una lettura più approfondita è possibile consultare la referenza [2].

3.5 Informazione classica e quantistica

La teoria dell'informazione quantistica presenta molte differenze con la teoria dell'informazione classica. Consideriamo che due persone, Alice e Bob. Supponiamo che Alice disponga di una sorgente classica che produce i simboli $X = 0, 1, ..., n$ con probabilità $p_0, ..., p_n$ e che Bob debba determinare il valore di X. Per fare ciò Alice prepara uno stato quantistico ρ_X scelto nel set $\rho_0, ..., \rho_n$ e lo spedisce a Bob che esegue una misura su tale stato e fa la sua ipotesi sul valore di X basata sul risultato della sua misura Y. Una misura dell'informazione che Bob ha acquisito può essere fornita dall'informazione mutua $H(X : Y)$. In generale vale $H(X : Y) \leq H(X)$. L'obiettivo di Bob è quello di massimizzare $H(X : Y)$ per far si che sia più vicina possibile a $H(X)$. Definiamo *l'informazione accessibile* a Bob il massimo di $H(X : Y)$ su tutte le possibili misure che può compiere. Classicamente l'informazione accessibile non ha grande importanza dal momento che due stati classici sono, almeno teoricamente, sempre distinguibili. Quantisticamente ciò non è possibile nemmeno in linea di principio: non esiste una procedura per distingue stati quantistici non ortogonali. Se Alice prepara gli stati non ortogonali $|\psi\rangle$ e $|\phi\rangle$ rispettivamente con probabilità p e $1 - p$ l'informazione accessibile è strettamente minore di $H(p)$. Al contrario se Alice ha preparato un bit classico nello stato 0 con probabilità p o nello stato 1 con probabilità $1 - p$ non c'è ragione per cui Bob non debba essere in grado di distinguere questi due stati per cui l'informazione accessibile è pari all'entropia della preparazione $H(p)$.

Un'altra differenza cruciale risiede nel *no-cloning theorem* che afferma che non è possibile creare un dispositivo che, ricevendo in ingresso uno stato $|\psi\rangle$ o uno stato $|\phi\rangle$ non ortogonali, possa creare una copia dello stato in ingresso, ovvero $|\psi\rangle|\psi\rangle$ o $|\phi\rangle|\phi\rangle$.

3.6 Limite di Holevo

Un metodo esplicito per calcolare l'informazione accessibile non esiste, ma esiste un teorema che fornisce un limite superiore per l'informazione accessibile detto *limite di Holevo*.
Supponiamo che Alice prepara uno stato ρ_X con $X = 0, 1..., n$ con probabilità $p_0, ..., p_n$. Bob esegue una misura descritta dagli elementi POVM $\{E_y\}$=$\{E_0, ..., E_m\}$ su quello stato e ottiene come risultato Y. Per qualunque misura di questo tipo vale

$$H(X : Y) \leq S(\rho) - \sum_x p_x S(\rho_x) \tag{3.25}$$

dove $\rho = \sum_x p_x \rho_x$. Il membro di sinistra viene spesso indicato con il simbolo χ.

Dimostrazione. La dimostrazione coinvolge tre sistemi quantistici: P, Q e M. Q è il sistema che Alice da' a Bob; P e M sono due sistemi fittizi ausiliari che semplificano la dimostrazione. Il sistema P può essere immaginato come il sistema di 'preparazione'. Per definizione esiste una base ortonormale $|x\rangle$ i cui elementi corrispondono alle etichette $0, ..., n$ delle possibili preparazioni del sistema Q. M può essere immaginato come l'apparato di misura di Bob e anche qui possiamo definire una base ortonormale $|y\rangle$ i cui elementi corrispondono ai possibili esiti della misura $1, ..., n$. Assumiamo che lo stato iniziale del sistema sia

$$\rho^{PQM} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x \otimes |o\rangle\langle 0| \tag{3.26}$$

dove il prodotto tensoriale è scritto con l'ordine PQM. Intuitivamente possiamo dire che Alice ha scelto un valore di x con probabilità p_x, ha prodotto lo stato ρ_x e lo ha dato a Bob che sta per effettuare una misura tramite il suo apparato che inizialmente si trova nello stato $|0\rangle$. Per descrivere la misura introduciamo un superoperatore Γ che agisce su Q e M, ma non su P. L'azione di Γ è quella di fare una misura con elementi POVM $\{E_y\}$ su Q e di memorizzare il risultato in M:

$$\Gamma(\sigma \otimes |0\rangle\langle 0|) \equiv \sum_y \sqrt{E_y}\sigma\sqrt{E_y} \otimes |y\rangle\langle y| \tag{3.27}$$

dove σ è lo stato di Q e $|0\rangle$ è lo stato iniziale del sistema di misura. È possibile mostrare che Γ preserva la traccia. Siccome M è non correlato con P e Q si ha $S(P : Q) = S(P : Q, M)$ e $S(P : Q, M) \geq S(P' : Q', M')$. Inoltre $S(P' : Q', M' \leq S(P' : M')$ per

quanto detto nella proprietà 8) dell'entropia di von Neumann. I sistemi primati indicano i sistemi dopo l'applicazione di Γ. Mettendo insieme questi risultati troviamo

$$S(P' : M') \leq S(P : Q). \tag{3.28}$$

Questa equazione è una espressione del limite di Holevo.

Calcoliamo infatti

$$\rho^{PQ} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x, \tag{3.29}$$

da cui si ha $S(P) = H(p_x),\ S(Q) = \rho$ e $S(P,Q) = H(p_x) + \sum_x p_x S(\rho_x)$ per la proprietà 10) dell'entropia di von Neumann. Troviamo quindi:

$$S(P : Q) = S(P) + S(Q) - S(P,Q) = S(\rho) - \sum_x p_x S(\rho_x) \tag{3.30}$$

che è il lato destro della disuguaglianza di Holevo. Per quanto riguarda il lato sinistro:

$$\rho^{P'Q'M'} = \sum_{x,y} p_x |x\rangle\langle x| \otimes \sqrt{E_y}\rho_x\sqrt{E_y} \otimes |y\rangle\langle y|. \tag{3.31}$$

Eseguendo la traccia parziale sul sistema Q' e osservando che la probabilità congiunta $p(x,y)$ per la coppia (X,Y) soddisfa $p(x,y) = p_x p(y|x) = p_x Tr[\rho_x E_y] = p_x Tr\left[\sqrt{E_y}\rho_x\sqrt{E_y}\right]$ troviamo:

$$\rho^{P'M'} = \sum_{x,y} p(x,y) |x\rangle\langle x| \otimes |y\rangle\langle y|. \tag{3.32}$$

Dal momento che $S(P' : M') = H(X : Y)$ abbiamo concluso la dimostrazione. □

Come conseguenza di questo teorema possiamo dimostrare quanto precedentemente affermato. La proprietà 10) afferma che $S(\rho) - \sum_x p_x S(\rho_x) \leq H(X)$ con l'uguaglianza che si verifica se e solo se gli stati ρ_x hanno supporti ortogonali. Se non hanno supporti ortogonali la disuguaglianza è stretta e ciò implica $H(X : Y) \leq H(X)$ e quindi risulta impossibile per Bob determinare con assoluta certezza X partendo dal risultato della misura, Y.

Bibliografia

[1] G. Benenti, G. Casati, G. Strini, *Principles of Quantum Computation and Quantum Information. Volume I: Basic Concepts*, World Scientific, 2004.

[2] Charles H. Bennet, *Notes on Landauer's principle, reversible computation, and Maxwell's Demon*, Elsevier Science Ltd., 2003.

[3] C. Degli Esposti Boschi, *Introduzione alla Teoria dell'Informazione Quantistica - Lezioni del corso di Teoria Quantistica della Materia*, 2018.

[4] Michael A. Nielsen, Isaac L. Chuang, *Quantum Information and Quantum Information*, Cambringe, 2000.

[5] Jhon Preskill, *Quantum Information and Computation*, California Institute of Technology, 1998.

[6] R. Zucchini, *Quantum Mechanic: Lezioni del corso di Meccanica Quantistica*, 2019.